(19)日本国特許庁（ＪＰ）　(12)**公開特許公報**（Ａ）　(11)特許出願公開番号
特開2002－63052
（P2002－63052A）
(43)公開日　平成14年２月28日(2002.2.28)

(51)Int.Cl.[7]	識別記号	FI		テーマコード(参考)
G06F 12/00	501	G06F 12/00	501A	5B045
15/177	682	15/177	682B	5B082

審査請求　未請求　請求項の数５　ＯＬ　（全 11 頁）

(21)出願番号　特願2000－246813(P2000－246813)

(22)出願日　平成12年８月16日(2000.8.16)

(71)出願人　000005223
富士通株式会社
神奈川県川崎市中原区上小田中４丁目１番１号

(72)発明者　中條　義久
愛知県名古屋市東区葵１丁目16番38号　株式会社富士通愛知エンジニアリング内

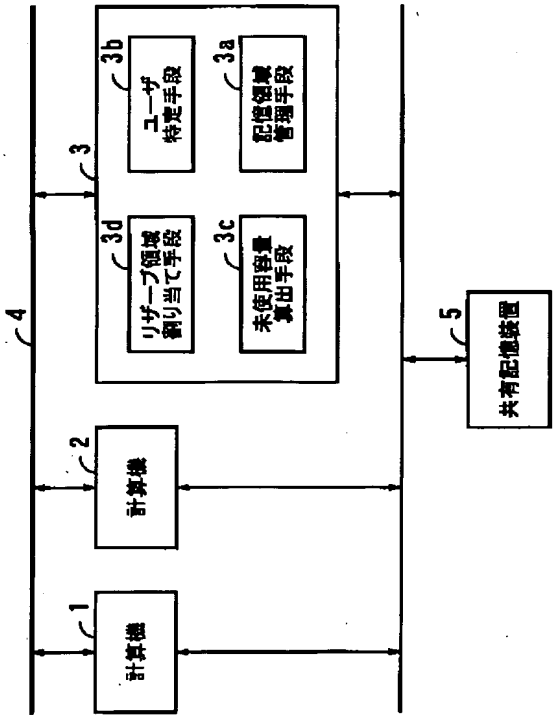
(72)発明者　矢口　聰彦
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

(74)代理人　100092152
弁理士　服部　毅巖

Ｆターム(参考)　5B045 DD03
5B082 CA01 CA08

(54)【発明の名称】　分散処理システム

(57)【要約】
【課題】　複数の計算機と、共有記憶装置とを有する分散処理システムにおいて、共有記憶装置への高速なアクセス性を実現する。
【解決手段】　記憶領域管理手段３ａは、各ユーザが使用可能な、共有記憶装置５上の記憶領域を管理する。ユーザ特定手段３ｂは、共有記憶装置５に対して書き込み要求を行ったユーザを特定する。未使用容量算出手段３ｃは、ユーザ特定手段３ｂによる特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する。リザーブ領域割り当て手段３ｄは、未使用容量算出手段３ｃによって算出された未使用容量に応じて、要求を行った計算機が独自に管理可能な共有記憶装置５上の領域であるリザーブ領域をその計算機に対して割り当てる。



1

【特許請求の範囲】

【請求項１】　相互に接続された複数の計算機と、少なくとも１つの共有記憶装置とから構成される分散処理システムにおいて、

各ユーザが使用可能な、前記共有記憶装置上の記憶領域を管理する記憶領域管理手段と、

前記共有記憶装置に対して書き込み要求を行ったユーザを特定するユーザ特定手段と、

前記ユーザ特定手段による特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する未使用容量算出手段と、

前記未使用容量算出手段によって算出された未使用容量に応じて、要求を行った計算機が独自に管理可能な前記共有記憶装置上の領域であるリザーブ領域をその計算機に対して割り当てるリザーブ領域割り当て手段と、

を有することを特徴とする分散処理システム。

【請求項２】　前記リザーブ領域割り当て手段は、未使用容量が所定量以上である場合は、未使用領域の一定量を要求を行った計算機に対して割り当て、

未使用容量が所定量を下回った場合には、システムを構成する計算機の台数で前記未使用容量を除して得られた値に対応するリザーブ領域を割り当てる、

ことを特徴とする請求項１記載の分散処理システム。

【請求項３】　前記リザーブ領域割り当て手段は、未使用容量が所定量を更に下回った場合には、リザーブ領域の割り当てを停止することを特徴とする請求項２記載の分散処理システム。

【請求項４】　前記記憶領域管理手段は、ユーザ群毎に使用可能な記憶領域を管理し、

前記未使用容量算出手段は、前記ユーザ特定手段によって特定されたユーザが所属するユーザ群に係る未使用容量を算出する、

ことを特徴とする請求項１記載の分散処理システム。

【請求項５】　相互に接続された複数の計算機と、少なくとも１つの共有記憶装置とから構成される分散処理方法において、

各ユーザが使用可能な、前記共有記憶装置上の記憶領域を管理する記憶領域管理ステップと、

前記共有記憶装置に対して書き込み要求を行ったユーザを特定するユーザ特定ステップと、

前記ユーザ特定手段による特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する未使用容量算出ステップと、

前記未使用容量算出手段によって算出された未使用容量に応じて、要求を行った計算機が独自に管理可能な前記共有記憶装置上の領域であるリザーブ領域をその計算機に対して割り当てるリザーブ領域割り当て手段ステップと、

を有することを特徴とする分散処理方法。

【発明の詳細な説明】

2

【０００１】

【発明の属する技術分野】本発明は分散処理システムに関し、特に、相互に接続された複数の計算機と、少なくとも１つの共有記憶装置とから構成される分散処理システムに関する。

【０００２】

【従来の技術】ＵＮＩＸ（登録商標）システムでは、複数台の計算機に処理を分散させる分散処理システムが提供されている。このような分散処理システムにおいては、各計算機によって処理されるデータがシステム全体として一貫性および整合性を保っていなければならないため、同じデータがシステム内に複数存在することは望ましいこととは言えない。従って、システム内にユニークに存在しているデータを格納する記憶装置は、各計算機から共通にアクセスできるようにする必要がある。このような目的のもと、一個もしくは物理的または論理的に分割された複数台の記憶装置（以下、「共有記憶装置」と称す）をシステム内の各計算機によって共有させる共有ファイルシステムが実現されている。

【０００３】このような共有ファイルシステムを複数のユーザが使用する場合には、一部のユーザが共有記憶装置を独占的に使用してしまって他のユーザが使用できないといった事態を未然に防ぐ必要が生ずる。このような事態を防止する対策として、システムの管理者が、予め各ユーザの使用量の上限を設定しておき、その上限を越えないように制御する方法が考えられる。また、ユーザ毎ではなく、例えば、部門単位や作業グループ単位といった、ひとまとまりのユーザ群ごとに使用量の上限を制御する方法が考えられる。

【０００４】ところで、従来は、前述のような記憶装置を複数台の計算機で共有するシステムは実現されておらず、類似する形態として、記憶装置を接続した計算機をサーバとし、他の計算機はサーバに対してファイルの操作を依頼するクライアントとして動作する形態が存在していた。このような、サーバ・クライアントシステムにおける、ユーザ毎またはユーザ群毎の記憶装置の使用量の上限に関する制御は次のように行われていた。

【０００５】図７は、従来のサーバ・クライアントシステムの構成例を示すブロック図である。この図において、計算機１０～１２は、通信路１３を介して相互に接続されている。計算機１０には、記憶装置１４が接続され、ファイルシステム１４ａが構築されている。計算機１０には、ファイルシステム１４ａのサービスを他の計算機に提供するためのサーバサブシステム１０ａが動作し、計算機１１，１２ではそれぞれクライアントサブシステム１１ａ，１２ａが動作している。

【０００６】ここで、計算機１１のユーザがファイルシステム１４ａにデータを格納する場合、このユーザはデータの書き込みをクライアントサブシステム１１ａに依頼する。クライアントサブシステム１１ａは、そのデー

タを通信路１３を介して計算機１０のサーバサブシステム１０ａに送信する。サーバサブシステム１０ａは、受信したデータを格納するための記憶装置の使用量と、予め設定されたそのユーザの使用量の上限を比較し、使用量の上限を越えない場合には、受信したデータをファイルシステム１４ａに供給して格納させる。

【０００７】

【発明が解決しようとする課題】しかし、このような方法では、計算機１１のユーザがファイルシステム１４ａ上に配置されたファイルにデータを書き込む度に、クライアントサブシステム１１ａがデータそのものを通信路１３を介してサーバサブシステム１０ａに送信する必要があるため、計算機１０およびサーバサブシステム１０ａの処理能力および通信路１３の伝送能力がデータを書き込む時の性能を決めてしまうという問題点があった。更に、計算機１１のユーザと、計算機１２のユーザが同じファイルシステム１４ａにデータを格納しようとした場合には処理の競合が発生し、性能の劣化が更に顕著となるという問題点もあった。

【０００８】前述の点に鑑み、分散処理システム全体として実行性能を向上させるために、共有記憶装置中のデータ記憶領域の管理をサーバから、各クライアントへ分権化することによって、各クライアントがサーバに対して記憶装置上のデータ書き込み対象のブロックを問い合わせることを不要にできるデータ管理方式が、特願平１１－１４３５０２号公報に開示されている。このデータ管理方式において提案された管理方式によると、各クライアントは、サーバから管理を委託された記憶領域（以下、「リザーブ領域」と称す）に関しては、サーバから独立して独自の裁量で管理することが可能となる。従って、共有記憶装置に対してデータを書き込む毎にサーバに問い合わせる必要もなく、高速なアクセスが可能となる。

【０００９】図８は、前述の共有記憶装置を用いた方式の構成例を示す図である。この図において、計算機２０～２２は、通信路２３を介して相互に接続されている。計算機２０～２２には、直接アクセス可能な共有記憶装置２５が接続され、共有ファイルシステム２５ａが構築されている。計算機２０には、共有ファイルシステム２５ａを管理するための管理用サブシステム２０ａが動作し、計算機２１，２２には、アクセス用サブシステム２１ａ，２２ａがそれぞれ動作している。

【００１０】ここで、計算機２１のユーザが共有ファイルシステム２５ａにデータを格納する場合を考えると、このユーザはデータの書き込みを、アクセス用サブシステム２１ａに依頼する。アクセス用サブシステム２１ａは、そのデータがアクセス用サブシステム２１ａが管理するリザーブ領域に格納できるものであれば、リザーブ領域内にそのデータを書き込むための領域を割り当て、共有記憶装置２５に直接書き込む。

【００１１】このときの使用量の上限に関する制御を考えると、共有ファイルシステム２５ａにおける使用量の上限は、ユーザが計算機２１および計算機２２の何れの計算機上で使用していたとしても、また、双方の計算機上で同時に使用していたとしても、総合的に判断または制御されなければならない。そのためには、管理用サブシステム２０ａがユーザ毎の使用量の上限を一括的に管理する、従来の手法を踏襲した方法が簡明な実現方式であると考えられる。

【００１２】しかしながら、このような従来方式を踏襲した方法では、計算機２１のユーザが共有ファイルシステム２５ａにデータを書き込む度に、アクセス用サブシステム２１ａは、管理用サブシステム２０ａに対して書き込みの可否の確認を行う必要がある。その結果、計算機２０および管理用サブシステム２０ａの処理能力ならびに通信路２３の伝送能力が、データを書き込む際の性能を決定してしまう。書き込むデータ自体の送信は不要であるので、図７の場合と比較すると影響は少ないが、データを書き込む毎に通信路２３を介して情報を送受信する必要があるため、特願平１１－１４３５０２号公報に開示されたデータ管理方式の特徴である高速アクセス性が阻害されてしまうという問題点もあった。

【００１３】ところで、近年では、分散処理システムの規模は拡大の一途を辿り、数百ないし数千の計算機を接続した分散処理システムも出現している。また、インターネットサービスプロバイダのように、多くのユーザに対して記憶装置を貸与するサービスも出現している。しかしながら、従来の手法を用いた分散処理システムでは、前述したように高速なアクセスが困難であることから、ユーザの要求を満足するシステムの構築が困難であるという問題点もあった。

【００１４】本発明は、このような点に鑑みてなされたものであり、特願平１１－１４３５０２号公報に開示されたデータアクセス管理方式によって実現された高速なアクセスを阻害することなく、また、分散処理システムに対するシステムの負荷を最小限に抑えて記憶装置の制限制御を行うことが可能な分散処理システムを提供することを目的とする。

【００１５】

【課題を解決するための手段】本発明では上記課題を解決するために、図１に示す、相互に接続された複数の計算機１～３と、少なくとも１つの共有記憶装置５とから構成される分散処理システムにおいて、各ユーザが使用可能な、共有記憶装置５上の記憶領域を管理する記憶領域管理手段３ａと、共有記憶装置５に対して書き込み要求を行ったユーザを特定するユーザ特定手段３ｂと、ユーザ特定手段３ｂによる特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する未使用容量算出手段３ｃと、未使用容量算出手段３ｃによって算出された未使用容量に応じて、要求を行った計算機が独自

に管理可能な共有記憶装置５上の領域であるリザーブ領域をその計算機に対して割り当てるリザーブ領域割り当て手段３ｄと、を有することを特徴とする分散処理システムが提供される。

【００１６】ここで、記憶領域管理手段３ａは、各ユーザが使用可能な、共有記憶装置５上の記憶領域を管理する。ユーザ特定手段３ｂは、共有記憶装置５に対して書き込み要求を行ったユーザを特定する。未使用容量算出手段３ｃは、ユーザ特定手段３ｂによる特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する。リザーブ領域割り当て手段３ｄは、未使用容量算出手段３ｃによって算出された未使用容量に応じて、要求を行った計算機が独自に管理可能な共有記憶装置５上の領域であるリザーブ領域をその計算機に対して割り当てる。

【００１７】

【発明の実施の形態】以下、本発明の実施の形態を図面を参照して説明する。図１は、本発明の動作原理を説明する原理図である。この図に示すように、計算機１～３は、通信路４によって相互に接続され、分散処理システムを構成している。また、計算機１～３のそれぞれは、共有記憶装置５に接続されており、必要なデータを書き込んだり、読み出したりすることが可能とされている。

【００１８】計算機３は、計算機１，２に対してリザーブ領域を供与する処理を行う。ここで、この図に示すように、計算機３は、記憶領域管理手段３ａ、ユーザ特定手段３ｂ、未使用容量算出手段３ｃ、および、リザーブ領域割り当て手段３ｄによって構成されている。

【００１９】記憶領域管理手段３ａは、各ユーザが使用可能な、共有記憶装置５上の記憶領域を管理する。ユーザ特定手段３ｂは、共有記憶装置５に対して書き込み要求を行ったユーザを特定する。

【００２０】未使用容量算出手段３ｃは、ユーザ特定手段３ｂによる特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する。リザーブ領域割り当て手段３ｄは、未使用容量算出手段３ｃによって算出された未使用容量に応じて、要求を行った計算機が独自に管理可能な共有記憶装置５上の領域であるリザーブ領域をその計算機に対して割り当てる。

【００２１】次に、以上の原理図の動作について説明する。いま、ユーザＡが、計算機１を操作して、共有記憶装置５に対して所定量のデータを書き込む要求を初めて行ったとする。すると、計算機１は、自己に割り当てられているリザーブ領域と、書き込もうとするデータ量とを比較し、書き込み可能である場合にはリザーブ領域に対して書き込みを実行する。リザーブ領域が不足している場合には、計算機３に対してリザーブ領域の確保を要請する。いま、最初の処理であるとし、計算機１にはリザーブ領域が供与されていないとすると、計算機１は計算機３に対してリザーブ領域の確保を要請する。

【００２２】要求を受けた計算機３では、ユーザ特定手段３ｂがどのユーザから要求がなされたかを特定し、特定結果を記憶領域管理手段３ａに通知する。記憶領域管理手段３ａは、各ユーザが使用可能な最大領域を示す情報を格納しており、この情報を未使用容量算出手段３ｃに対して供給する。

【００２３】未使用容量算出手段３ｃは、そのユーザが使用可能な領域のうち、未使用となっている領域の容量を算出し、リザーブ領域割り当て手段３ｄに通知する。リザーブ領域割り当て手段３ｄは、未使用容量算出手段３ｃから通知された未使用容量に応じて、適切な容量のリザーブ領域を計算機１に対して割り当てる。ここで、適切な容量とは、例えば、このユーザの記憶領域の使用量が最大領域の５０％未満である場合には、未使用容量の２５％をリザーブ領域として供与する。

【００２４】また、使用量が最大領域の５０％以上かつ９０％未満である場合には、未使用容量をシステムを構成する計算機の数（この例では“３”）で除した値に対応する領域をリザーブ領域として供与する。

【００２５】更に、使用量が最大領域の９０％以上である場合には、リザーブ領域は供与しない。なお、このようにして適切なリザーブ領域を、未使用容量から決定するのは、このユーザＡが、例えば、計算機２から、リザーブ領域を確保する要求を行った場合に、未使用領域が少ない場合には、例えば、計算機１からリザーブ領域を返還させる必要が生じ、その処理にリザーブ領域を確保する以上に長大な時間を要するからである。

【００２６】このようにして、適切な容量のリザーブ領域が計算機１に対して供与されると、計算機１は、このリザーブ領域を独自の裁量で管理し、ユーザＡから要求があったデータを書き込む。

【００２７】以上のように、計算機３に対してリザーブ領域が要求された場合には、そのユーザに係る未使用容量に応じて、適切なリザーブ領域を供与するようにしたので、ユーザ毎に使用制限を行うとともに、高速なアクセスを実現することが可能となる。

【００２８】次に、本発明の実施の形態の構成例について説明する。図２は、本発明の実施の形態の構成例を示す図である。この図に示すように、計算機４０～４２は、通信路４３によって相互に接続され、分散処理システムが構築されている。また、共有記憶装置４５は、通信路４４によって計算機４０～４２に接続されている。

【００２９】計算機４１～４２は、例えば、パーソナルコンピュータによって構成されている。通信路４３は、例えば、インターネットによって構成されている。通信路４４は、例えば、ＬＡＮ（Local Area Network）によって構成されている。

【００３０】計算機４０は、共有ファイルシステム４５ａを管理するための管理用サブシステム４０ａを有している。また、計算機４１，４２は、共有記憶装置４５に

アクセスするためのアクセス用サブシステム４１ａ，４２ａをそれぞれ有している。

【００３１】共有記憶装置４５は、例えば、ハードディスクによって構成されており、共有ファイルシステム４５ａが構築されている。また、共有ファイルシステム４５ａには、各ユーザのユーザＩＤと、使用可能な最大ブロック数等を示す情報が格納されている。

【００３２】次に、以上の実施の形態の動作について説明する。先ず、具体的な動作の説明にはいる前に、システム全体の動作の概略について説明する。共有ファイルシステム４５ａの管理テーブル４５ｂには、図３に示す情報が格納されている。ここで、グループＩＤは、各グループに付与されたユニークな番号である。使用量は、各グループが使用可能な共有ファイルシステム４５ａの記憶容量である。ユーザＩＤは、各グループを構成するユーザのＩＤである。最大使用ブロック数は、各ユーザが使用可能な最大のブロック数である。最大使用ファイル数は、各ユーザが使用可能な最大のファイル数である。なお、グループに属する全てのユーザの使用ブロック数の合計は、グループの使用量と等しい関係にある。

【００３３】この例では、グループＩＤが「Ｇ０００１」に属するユーザ群に対しては、５５ＧＢの記憶領域が割り当てられており、また、グループＩＤが「Ｇ０００２」に属するユーザ群に対しては、８５ＧＢの記憶領域が割り当てられている。グループＩＤが「Ｇ０００１」のグループには、ユーザＩＤが「Ｐ１００１～Ｐ１１００」のユーザが属しており、各ユーザが使用可能な最大ブロック数と、最大使用ファイル数とが示されている。また、グループＩＤが「Ｇ０００２」のグループには、ユーザＩＤが「Ｐ２００１～Ｐ２１００」のユーザが属しており、前述の場合と同様に各ユーザが使用可能な最大ブロック数と、最大使用ファイル数とが示されている。このように、各ユーザは、使用可能なブロック数とファイル数とが予め決定されており、その決定された数量を上回って使用することはできない。

【００３４】管理用サブシステム４０ａは、各ユーザ単位で現在の記憶領域の使用状況を管理する。各ユーザに割り当てられた記憶領域は、図４に示すように、使用済みブロック、リザーブ領域、および、未使用ブロックから構成されている。ここで、使用済みブロックは、既に使用されているブロックを示す。リザーブ領域は、各計算機が独自の裁量で使用可能な領域である。未使用ブロックは、未だ使用されていないブロックを示す。各計算機は、リザーブ領域内で処理が可能な場合には、独自の裁量でリザーブ領域を使用することにより、ユーザからの要求に応える。そして、リザーブ領域が不足した場合や、無くなった場合には、管理用サブシステム４０ａに対してリザーブ領域の供与を要求する。管理用サブシステム４０ａは、その時点における未使用ブロック数を参照し、その容量に応じて最適なリザーブ領域を供与する。具体的には、管理用サブシステム４０ａは、以下の処理に従って、要求を行った計算機に対してリザーブ領域を供与する。

（１）未使用ブロック数が、最大使用ブロック数の５０％を上回っている場合には、未使用ブロックの２５％をリザーブ領域として要求を行った計算機に供与する。

（２）未使用ブロック数が最大使用ブロック数の５０％以下かつ１０％以上であり、また、未使用ブロック数が５０メガバイトを下回らない場合には、管理用サブシステム４０ａは、未使用ブロックの容量をシステムを構成する計算機の台数で除して得られた値に対応する領域を、リザーブ領域として要求を行った計算機に供与する。

（３）未使用ブロック数が最大使用ブロック数の１０％未満であり、かつ、未使用ブロック数の容量が５０メガバイトを下回った場合には、管理用サブシステム４０ａは、リザーブ領域の供与を行わない。その結果、アクセス用サブシステム４１ａ，４２ａは、書き込み要求が発生するたびに管理用サブシステム４０ａに対して書き込みの可否の問い合わせを行い、許可が得られた場合には書き込み処理を実行する。

【００３５】このように、未使用ブロック数に応じてリザーブ領域の割り当て量を変化させることにより、高速なアクセスを可能とすることができる。以下にその動作の詳細について説明する。

【００３６】いま、計算機４１のユーザＡが、１０００ブロックからなるデータの共有ファイルシステム４５ａへの書き込みを、システムに対して初めて要求したとする。なお、ユーザＡは、他の計算機からも未だ書き込み要求を行っていないものとする。従って、ユーザＡの記憶領域は、全て未使用状態である。

【００３７】すると、計算機４１のアクセス用サブシステム４１ａは、書き込み要求の１０００ブロックと、現在有しているリザーブ領域（＝０）とを比較し、リザーブ領域が不足しているので、管理用サブシステム４０ａに対して、リザーブ領域を供与するように要求する。

【００３８】管理用サブシステム４０ａは、ユーザＡが記憶領域をまだ使用していないことを検知し、前述の（１）に該当することから、未使用ブロックの２５％を計算機４１に対してリザーブ領域として供与する。具体的には、ユーザＡの最大使用ブロックが１００００ブロックである場合には、２５００（＝１００００×０．２５）ブロックがリザーブ領域として計算機４１に対して供与される。

【００３９】その結果、計算機４１のアクセス用サブシステム４１ａは、２５００ブロックのリザーブ領域を確保し、そのうちの１０００ブロックを要求されたデータの書き込みに使用する。

【００４０】次に、ユーザＡが計算機４２から、１１００ブロックからなるデータを共有ファイルシステム４５

ａに対して書き込む要求を行ったとすると、前述の場合と同様の処理が実行され、計算機４２のアクセス用サブシステム４２ａに対して１８７５（７５００×０．２５）ブロックのリザーブ領域が供与される。アクセス用サブシステム４２ａは、供与された１８７５ブロックのうち、１１００ブロックを要求されたデータの書き込みに対して割り当てる。

【００４１】続いて、ユーザＡが計算機４２から、５００ブロックからなるデータを共有ファイルシステム４５ａに対して書き込む要求を行ったとすると、その時点での計算機４２が有するユーザＡのリザーブ領域の残量は７７５ブロックであるので、アクセス用サブシステム４２ａは、リザーブ領域のうち５００ブロックを要求されたデータの書き込みに割り当てる。

【００４２】続いて、ユーザＡが計算機４２から、１１００ブロックからなるデータを共有ファイルシステム４５ａに対して書き込む要求を行ったとすると、その時点での計算機４２が有するユーザＡのリザーブ領域の残量は２７５ブロックであるので、アクセス用サブシステム４２ａは、リザーブ領域の供与を管理用サブシステム４０ａに対して要求する。

【００４３】このとき、ユーザＡの未使用ブロックは、５６２５ブロックであり全体の５０％以上であることから、前述の場合と同様にその２５％に該当する１４０６ブロックが計算機４２に対して割り当てられることになる。その結果、未使用領域は４２１９（＝５６２５－１４０６）となり、また、計算機４２のリザーブ領域は、１６８１（＝２７５＋１４０６）ブロックとなる。計算機４２は、リザーブ領域のうち１１００ブロックを、要求されたデータの書き込みに割り当てる。

【００４４】更に、ユーザＡが、計算機４２から、１０００ブロックからなるデータを共有ファイルシステム４５ａに対して書き込む要求を行ったとすると、その時点での計算機４２が有するユーザＡのリザーブ領域の残量は５８１ブロックであるので、アクセス用サブシステム４２ａは、リザーブ領域の供与を管理用サブシステム４０ａに対して要求する。ここで、ユーザＡの未使用ブロックは、４２１９ブロックであり、全体の５０％を下回っているので、管理用サブシステム４０ａは、前述の（２）の処理により、リザーブ領域を割り当てる。即ち、その時点における未使用領域の４２１９ブロックを、システムを構成する計算機の台数である３で除して得られた値に対応する１４０６ブロックをリザーブ領域として計算機４２に供与する。その結果、計算機４２は、供与された１４０６ブロックのうち、１０００ブロックを要求されたデータの書き込みに割り当てる。

【００４５】以上のような処理が繰り返され、ユーザＡの未使用ブロックが１０００ブロックを下回った場合において、更に、リザーブ領域の要求がなされた場合には、前述の（３）の処理が実行される。従って、これ以降は、リザーブ領域は供与されないので、書き込み要求を受けた計算機が、管理サブシステム４０ａに対して書き込みの可否を直接問い合わせ、許可された場合には要求されたデータの書き込み処理を実行することになる。

【００４６】以上に示したように、本実施の形態によれば、ユーザ単位で使用可能な領域である最大ブロック数を定義するとともに、使用可能な領域の残量に応じて、リザーブ領域を供与するようにしたので、例えば、所定量を一律にリザーブ領域として供与する場合と比較すると、未使用ブロックに応じた最適なリザーブ領域を供与することができる。具体的には、未使用ブロックが少なくなった場合に対応する（２）の処理では、未使用ブロックをシステムを構成する計算機の台数で除した値に対応する領域をリザーブ領域として供与することにより、各計算機が一定量のリザーブ領域を確実に確保することが可能となり、リザーブ領域が不足することに起因して、管理用サブシステム４０ａに対してリザーブ領域の要求が頻繁に行われることを防止できる。

【００４７】また、未使用ブロックがかなり少なくなった場合に対応する（３）の処理では、管理用サブシステム４０ａがリザーブ領域を供与することを停止することにより、リザーブ領域が不足することによって発生する他の計算機からのリザーブ領域の回収処理の発生を防止し、処理速度を向上させることが可能となる。具体的には、例えば、計算機４１が２０００ブロックのリザーブ領域を有しており、現時点における未使用ブロックが５００ブロックである場合に、１５００ブロックのリザーブ領域が計算機４２から要求された場合には、計算機４１から不足分のブロックを回収する処理が必要となるが、本実施の形態によれば、このような処理の発生を防止できる。

【００４８】最後に、図５，６を参照して、以上に説明した処理を可能にするためのフローチャートについて説明する。図５は、書き込み要求がなされたアクセス用サブシステムにおいて実行される処理の一例を説明するフローチャートである。このフローチャートが開始されると、以下の処理が実行される。

【００４９】［Ｓ１０］アクセス用サブシステムは、計算機から書き込み要求を受信したか否かを判定し、書き込み要求を受信した場合にはステップＳ１１に進み、それ以外の場合には同一の処理を繰り返す。

［Ｓ１１］アクセス用サブシステムは、書き込みの要求を行ったユーザを特定する。

【００５０】［Ｓ１２］アクセス用サブシステムは、ステップＳ１１において特定したユーザに対応するリザーブ量Ｒａを取得する。ここで、リザーブ量Ｒａとは、要求を行ったユーザが現在使用している計算機に供与されているリザーブ領域の容量である。

［Ｓ１３］アクセス用サブシステムは、ユーザが行った書き込み要求の要求ブロック数Ｗａを取得する。

【００５１】［Ｓ１４］アクセス用サブシステムは、要求ブロック数Ｗａがリザーブ量Ｒａよりも大きいか否かを判定し、大きい場合にはステップＳ１５に進み、それ以外の場合にはステップＳ１７に進む。

［Ｓ１５］アクセス用サブシステムは、管理用サブシステム４０ａに問い合わせを行って、リザーブ領域を獲得する処理を実行する。なお、この処理の詳細は、図６を参照して説明する。

【００５２】［Ｓ１６］アクセス用サブシステムは、既存のリザーブ量Ｒａに対して、新たに獲得した新規リザーブ量を加算し、リザーブ量Ｒａとする。

［Ｓ１７］アクセス用サブシステムは、確保したリザーブ領域を利用して書き込み処理を実行する。

［Ｓ１８］アクセス用サブシステムは、リザーブ量Ｒａから、書き込みによって使用した領域に対応するＷａを減算し、現在のリザーブ量Ｒａを算出する。

【００５３】以上の処理によれば、計算機において書き込み要求がなされた場合には、必要に応じて管理用サブシステム４０ａに対してリザーブ領域の確保を要請し、要求されたデータを共有ファイルシステム４５ａに書き込むことが可能となる。

【００５４】次に、図６を参照して、図５に示すリザーブ領域獲得処理の詳細について説明する。このフローチャートが開始されると、以下の処理が実行される。

［Ｓ３０］管理用サブシステム４０ａは、変数Ｃａに対して、このユーザの使用済みブロック数を代入する。

【００５５】［Ｓ３１］管理用サブシステム４０ａは、変数Ｌａに対して、このユーザの最大使用ブロック数を代入する。

［Ｓ３２］管理用サブシステム４０ａは、変数Ｎａに対して、新たに獲得する新規リザーブ量を代入する。

【００５６】［Ｓ３３］管理用サブシステム４０ａは、変数Ｎｍに対して、ファイルシステムを共有する計算機の台数を代入する。

［Ｓ３４］管理用サブシステム４０ａは、使用済みブロック数Ｃａの値が、最大使用ブロック数Ｌａに０．５を乗算した値以下である場合には、ステップＳ３９に進み、それ以外の場合にはステップＳ３５に進む。

【００５７】［Ｓ３５］管理用サブシステム４０ａは、使用済みブロック数Ｃａの値が、最大使用ブロック数に０．９を乗算した値よりも大きい場合にはステップＳ３７に進む、それ以外の場合にはステップＳ３６に進む。

［Ｓ３６］管理用サブシステム４０ａは、最大使用ブロック数Ｌａから使用済みブロック数Ｃａを減算した値が５０ＭＢ未満である場合にはステップＳ３７に進み、それ以外の場合にはステップＳ３８に進む。

【００５８】［Ｓ３７］管理用サブシステム４０ａは、新規リザーブ量Ｎａに０を代入し、ステップＳ４０に進む。即ち、新たにリザーブ領域を付与しないとして次の処理に進む。

【００５９】なお、この処理は、前述の（３）の処理に対応している。

［Ｓ３８］管理用サブシステム４０ａは、最大使用ブロック数Ｌａから使用済みブロック数Ｃａを減算して得られた値を計算機の台数Ｎｍで除算し、得られた値を新規リザーブ量Ｎａに代入し、ステップＳ４０に進む。

【００６０】なお、この処理は、前述の（２）の処理に対応している。

［Ｓ３９］管理用サブシステム４０ａは、最大使用ブロック数Ｌａから使用済みブロック数Ｃａを減算して得られた値に、０．２５を乗算した値を、新規リザーブ量Ｎａに代入し、ステップＳ４０に進む。

【００６１】なお、この処理は、前述の（１）の処理に対応している。

［Ｓ４０］管理用サブシステム４０ａは、新規のリザーブ領域をＮａだけ確保して、要求を行った計算機に対して供与する。そして、もとの処理に復帰（リターン）する。

【００６２】以上に示すフローチャートによれば、前述した実施の形態において示す機能を実現することが可能となる。なお、以上の実施の形態において示したリザーブ領域を確保する方法は、あくまでも一例であって、これ以外にも種々の実現形態が考えられる。要は、各ユーザに割り当てられた最大使用ブロックの残量に応じて、割り当て方法を変更するようにすれば、本発明の目的は達成されるものと考えられる。

【００６３】また、以上の実施の形態では、最大使用ブロックを基準にして、最適なリザーブ領域を計算機に供与するようにしたが、最大使用ファイル数を基準として同様の判断を行うことも可能である。

【００６４】更に、これまで述べてきたように、リザーブ領域やリザーブ量の概念を取り入れると、ディスク使用量の制限管理を無効の状態から有効の状態に変更する際に、特別の考慮が必要となる。制限管理を無効化している状態において、現在の使用量を管理しないこととすると、制限管理を有効化した際にファイルシステム内の全てのファイルについて走査し、使用者毎の使用状況を集計する必要があり、大量のデータを扱う分散処理システムにおいては現実的ではない。従って、制限管理を無効化している状態においても現在の使用量は情報を更新する方式を採る。

【００６５】ディスク使用量の制限管理を無効の状態から有効の状態に変更する場合、管理用サブシステムは、全てのアクセス用サブシステムに対してディスク使用量の制限管理を開始することを通知する。この通知を受けたアクセス用サブシステムは、以降初めて管理用サブシステムに対して割り当てを要求するに先立って、その時点において管理用サブシステムに未通知分の使用量を通知する。これによって、ディスク使用量の制限管理を正確かつ効率的に有効の状態に変更することが可能とな

る。

【００６６】最後に、上記の処理機能は、コンピュータによって実現することができる。その場合、分散処理システムが有すべき機能の処理内容は、コンピュータで読み取り可能な記録媒体に記録されたプログラムに記述されており、このプログラムをコンピュータで実行することにより、上記処理がコンピュータで実現される。コンピュータで読み取り可能な記録媒体としては、磁気記録装置や半導体メモリ等がある。市場へ流通させる場合には、ＣＤ－ＲＯＭ(Compact Disk Read Only Memory)やフロッピー（登録商標）ディスク等の可搬型記録媒体にプログラムを格納して流通させたり、ネットワークを介して接続されたコンピュータの記憶装置に格納しておき、ネットワークを通じて他のコンピュータに転送することもできる。コンピュータで実行する際には、コンピュータ内のハードディスク装置等にプログラムを格納しておき、メインメモリにロードして実行する。

【００６７】（付記１）　相互に接続された複数の計算機と、少なくとも１つの共有記憶装置とから構成される分散処理システムにおいて、各ユーザが使用可能な、前記共有記憶装置上の記憶領域を管理する記憶領域管理手段と、前記共有記憶装置に対して書き込み要求を行ったユーザを特定するユーザ特定手段と、前記ユーザ特定手段による特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する未使用容量算出手段と、前記未使用容量算出手段によって算出された未使用容量に応じて、要求を行った計算機が独自に管理可能な前記共有記憶装置上の領域であるリザーブ領域をその計算機に対して割り当てるリザーブ領域割り当て手段と、を有することを特徴とする分散処理システム。

【００６８】（付記２）　前記リザーブ領域割り当て手段は、未使用容量が所定量以上である場合は、未使用領域の一定量を要求を行った計算機に対して割り当て、未使用容量が所定量を下回った場合には、システムを構成する計算機の台数で前記未使用容量を除して得られた値に対応するリザーブ領域を割り当てる、ことを特徴とする付記１記載の分散処理システム。

【００６９】（付記３）　前記リザーブ領域割り当て手段は、未使用容量が所定量を更に下回った場合には、リザーブ領域の割り当てを停止することを特徴とする付記２記載の分散処理システム。

【００７０】（付記４）　前記記憶領域管理手段は、ユーザ群毎に使用可能な記憶領域を管理し、前記未使用容量算出手段は、前記ユーザ特定手段によって特定されたユーザが所属するユーザ群に係る未使用容量を算出する、ことを特徴とする付記１記載の分散処理システム。

【００７１】（付記５）　前記記憶領域管理手段による管理を無効状態から有効状態に変更する場合には、各計算機はその時点における記憶領域の使用量であって、未通知の使用量について前記記憶領域管理手段に通知することを特徴とする付記１記載の分散処理システム。

【００７２】（付記６）　相互に接続された複数の計算機と、少なくとも１つの共有記憶装置とから構成される分散処理方法において、各ユーザが使用可能な、前記共有記憶装置上の記憶領域を管理する記憶領域管理ステップと、前記共有記憶装置に対して書き込み要求を行ったユーザを特定するユーザ特定ステップと、前記ユーザ特定手段による特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する未使用容量算出ステップと、前記未使用容量算出手段によって算出された未使用容量に応じて、要求を行った計算機が独自に管理可能な前記共有記憶装置上の領域であるリザーブ領域をその計算機に対して割り当てるリザーブ領域割り当て手段ステップと、を有することを特徴とする分散処理方法。

【００７３】

【発明の効果】以上説明したように本発明では、相互に接続された複数の計算機と、少なくとも１つの共有記憶装置とから構成される分散処理システムにおいて、各ユーザが使用可能な、共有記憶装置上の記憶領域を管理する記憶領域管理手段と、共有記憶装置に対して書き込み要求を行ったユーザを特定するユーザ特定手段と、ユーザ特定手段による特定結果に応じて、要求を行ったユーザに係る未使用領域の容量を算出する未使用容量算出手段と、未使用容量算出手段によって算出された未使用容量に応じて、要求を行った計算機が独自に管理可能な共有記憶装置上の領域であるリザーブ領域をその計算機に対して割り当てるリザーブ領域割り当て手段と、を設けるようにしたので、各計算機に対して最適なリザーブ領域を割り当てることが可能となるので、高速なアクセス性を実現することが可能となる。

【図面の簡単な説明】

【図１】本発明の動作原理を説明する原理図である。

【図２】本発明の実施の形態の構成例を示す図である。

【図３】図２に示す管理テーブルの一例を示す図である。

【図４】各ユーザに割り当てられた記憶領域の分割の態様の一例を示す図である。

【図５】図２に示すアクセス用サブシステムで実行される処理の一例を説明するフローチャートである。

【図６】図２に示す管理用サブシステムで実行される処理の一例を説明するフローチャートである。

【図７】従来の分散処理システムの構成例を示す図である。

【図８】従来の分散処理システムの他の構成例を示す図である。

【符号の説明】

１～３　計算機

３ａ　記憶領域管理手段

３ｂ　ユーザ特定手段

３ｃ　未使用領域算出手段

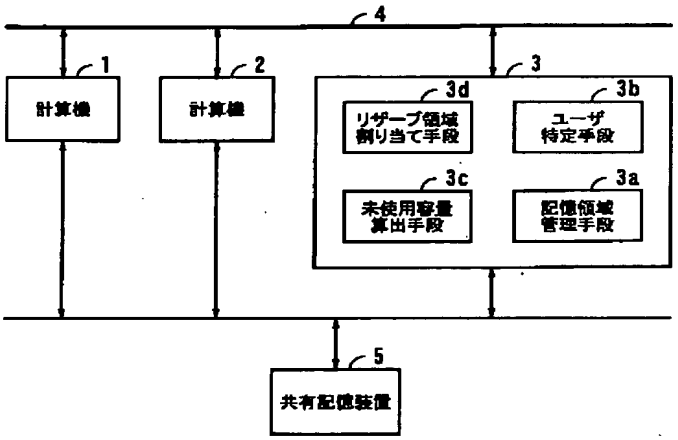
15

３ｄ　リザーブ領域割り当て手段
４　通信路
５　共有記憶装置
１０～１２　計算機
１０ａ　サーバサブシステム
１１ａ，１２ａ　クライアントサブシステム
１３　通信路
１４　記憶装置
１４ａ　ファイルシステム
２０～２２　計算機
２０ａ　管理用サブシステム

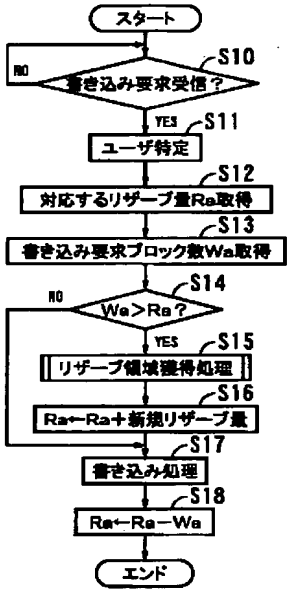
16

２１ａ，２２ａ　アクセス用サブシステム
２５　共有記憶装置
２５ａ　共有ファイルシステム
４０～４２　計算機
４０ａ　管理用サブシステム
４１ａ，４２ａ　アクセス用サブシステム
４３，４４　通信路
４５　共有記憶装置
４５ａ　共有ファイルシステム
４５ｂ　管理テーブル

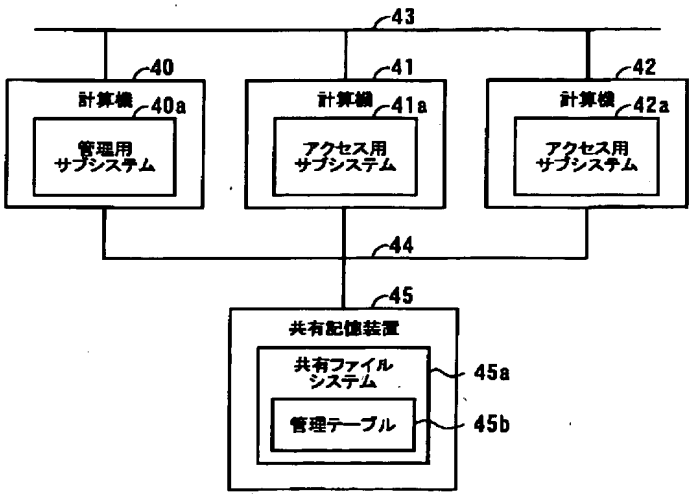
【図１】



【図５】



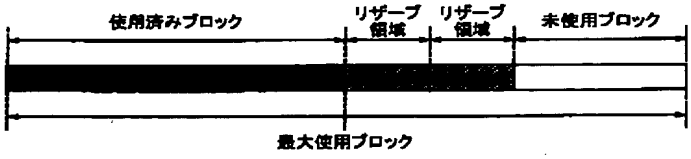
【図２】



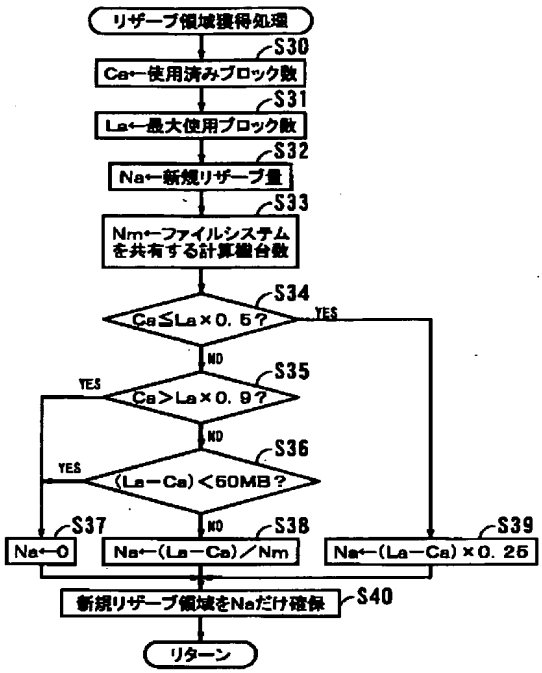
【図３】

グループID	使用量	ユーザID	最大使用ブロック数	最大使用ファイル数
G0001	55GB	P1001	12000	4000
		P1002	8000	3000
		：	：	：
		P1100	14000	6000
G0002	85GB	P2001	13000	5000
		P2002	8000	3000
		：	：	：
		P2100	10000	3500

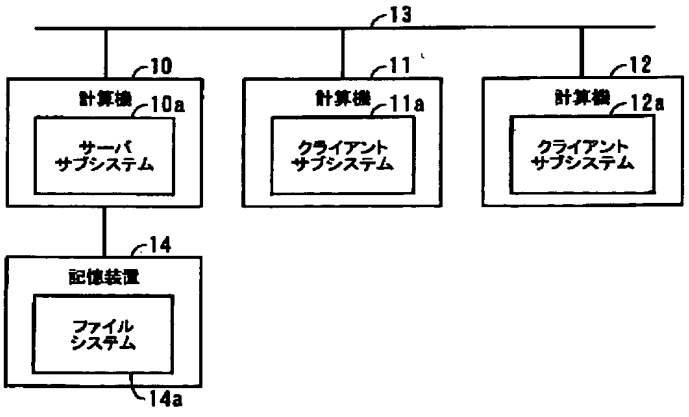
【図４】



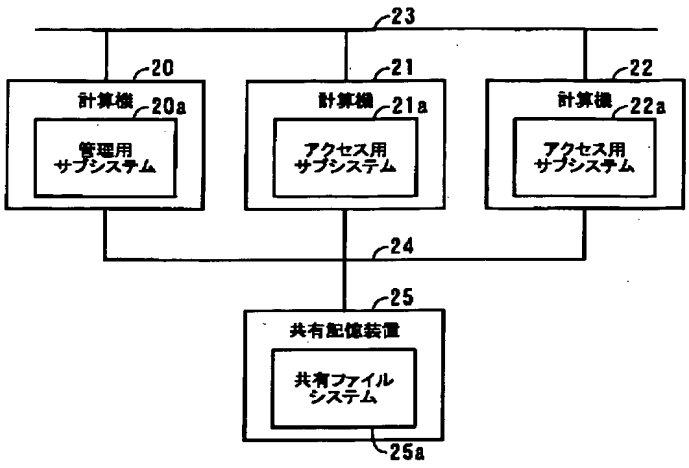
【図６】



【図７】



【図８】


23
20
計算機
20a
管理用
サブシステム
21
計算機
21a
アクセス用
サブシステム
22
計算機
22a
アクセス用
サブシステム
24
25
共有記憶装置
共有ファイル
システム
25a